THE STRUCTURE OF VARIATION AND ITS INFLUENCE ON THE ESTIMATION OF STATUS: INDICATORS OF CONDITION OF LAKES IN THE NORTHEAST, U.S.A.

THOMAS M. KINCAID[1]*, DAVID P. LARSEN[2] and N. SCOTT URQUHART[3]

[1] *Dynamac Corporation, Corvallis, Oregon, U.S.A.;* [2] *Western Ecology Division, National Health and Environmental Effects Research Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Corvallis, Oregon, U.S.A.;* [3] *Department of Statistics, Colorado State University, Fort Collins, Colorado, U.S.A.*

(author for correspondence, e-mail: kincaid.tom@epa.gov)*

(Received 15 October 2002; accepted 15 September 2003)

Abstract. One goal of regional-scale sample surveys is to estimate the status of a resource of interest from a statistically drawn representative sample of that resource. An expression of status is the frequency distribution of indicator scores capturing variability of attributes of interest. However, extraneous variability interferes with the status description by introducing bias into the frequency distributions. To examine this issue, we used data from a regional survey of lakes in the Northeast U.S. collected by the U.S. Environmental Protection Agency's Environmental Monitoring and Assessment Program (EMAP). We employ a components of variance model to identify sources of extraneous variance pertinent to status descriptions of physical, chemical, and biological attributes of the population of lakes in the NE. We summarize the relative magnitude of four components of variance (lake-to-lake, year, interaction, and residual) for each indicator and illustrate how extraneous variance biases the status descriptions. We describe a procedure that removes this bias from the status descriptions to produce unbiased estimates and introduce a novel method for estimating the 'cost' of removing the bias (expressed as either increased sampling uncertainty or additional samples needed to achieve the target precision in the absence of bias). We compare the relative magnitude of the four variance components across the array of indicators, finding in general that conservative chemical indicators are least affected by extraneous variance, followed by some nonconservative indicators, with nutrient indicators most affected by extraneous variance. Intermediate were trophic condition indicators (including sediment diatoms), fish species richness and individuals indicators, and zooplankton taxa richness and individuals indicators. We found no clear patterns in the relative magnitude of variance components as a function of several methods of aggregating fish and zooplankton indicators (e.g., level of taxonomy, or species richness vs. numbers of individuals).

Keywords: components of variance, lake condition indicators, status estimation

1. Introduction

One goal of surveys of ecological resources is to produce a snapshot description of the population or natural resource type of interest. A well designed and skillfully implemented survey will produce an accurate representation of the status

Environmental Monitoring and Assessment **98:** 1–21, 2004.

of the resource at the time the survey was conducted. The design of surveys, the measurement process, and some components of natural variation affect our ability to describe status by introducing extraneous variation and/or biases, producing a poorer representation of the population than may be desirable.

Urquhart *et al*. (1998) and Larsen *et al*. (2001) described a variance framework by which to organize the important sources of variation that affect estimation of status and detection of trends in regional surveys over time. This framework identified four major variance components: (1) population variation: site-to-site, spatial, or geographic variation in the particular attribute of interest, (2) coherent temporal (year) variation: year-to-year variation that affects all sites equally, (3) site-by-year interaction variation: year-to-year variation at an individual site that is not accounted for by coherent temporal variation, and (4) residual variation: the remaining variation not accounted for by the three other components, which includes team-to-team error, temporal variation within the sampling period, measurement error, analytical error, and errors in data processing. For further discussion regarding the constituents of residual variance see Larsen *et al*. (1995). This variance framework is similar to one proposed by Lewis (1978) and Matthews (1990) but identifies a coherent temporal variance component. Status descriptions intend to capture the frequency distribution of population variation, i.e., the frequency distribution of the site-to-site variation of an attribute of interest across the ecological resource. Status descriptions are affected by the relative magnitude of residual variation and site-by-year interaction variation and the absolute magnitude of coherent temporal variation. Trend detection, conversely, is affected by the absolute magnitude of the variance components. The effects on status descriptions and trend detection of all of the components except coherent temporal variation, however, can be managed through design and sample size choices. If coherent temporal variation is large, no amount of design manipulation can overcome its effect on regional trend detection. The only recourse is to wait a longer period of time than otherwise or to determine what controls this component so that its effect might be removed or minimized (see Urquhart *et al*., 1998).

Regional surveys of ecological condition often include measurement of an array of attributes that focus on endpoints of interest, such as the biological assemblages, as well as a variety of ancillary attributes that can be used to classify the resource types or, in combination with biological assemblage information, can be employed in associational analyses to help explain patterns in the endpoints of interest. If properly designed, these regional surveys can provide data by which to estimate the four variance components. In part in response to various calls for regional and national surveys of ecological resources, the U.S. EPA designed the Environmental Monitoring and Assessment Program (EMAP) with general goals to describe the status of ecological resources, their trends over time, and to assess probable cause of good or poor condition (Messer *et al*., 1991; Overton *et al*., 1991; Stevens, 1994). Other national and regional surveys of ecological resources include the National Resources Inventory (NRI) conducted by the U.S. Natural Resource Conservation Service, the Forest Inventory and Analysis (FIA) program conducted by the U.S. Forest Service, surveys conducted by the National Agricultural Statistical Service (NASS), the National Wetlands Inventory conducted by the U.S. Fish and Wildlife Service, and the North American Breeding Bird Survey sponsored by the U.S. National Biological Service (Olsen *et al*., 1999). Some of these surveys characterize ecological condition of these resources or have some components that do, while other surveys focus on harvestable crops produced by the ecosystems being surveyed.

As part of EMAP, a survey of lakes in the NE region of the U.S. covering the New England states and New York and New Jersey was designed and conducted (Larsen *et al*., 1994). This survey will be referenced as the Northeast Lakes Survey. The survey design allowed us to estimate the four major variance components identified above. In this paper, we present variance component estimates for the array of lake indicators sampled as part of the Northeast Lakes Survey, compare the spatial and temporal patterns of variability among the indicators, and consider the effect of taxonomic hierarchy on the variability of zooplankton and fish indicators. We discuss the influence of extraneous variation on status estimation and consider the relative performance of the indicator groups for status estimation. We describe a procedure that removes the bias introduced by extraneous variation and introduce a novel approach to estimate the cost (in terms of increased uncertainty of the status estimate, or in terms of the additional number of lakes that would have to be sampled to achieve the same precision as if extraneous variance was negligible) of removing the bias. We also examine relationships among the variance components across the set of indicators to address questions such as: What insight can we gain from the variance structure about improving field protocols for reducing extraneous variance? Are the relative magnitudes of the variance components related to the biological or geochemical activity of the attribute, e.g., is the variability structure of inactive chemical properties substantially different from that of reactive chemical properties? Are the patterns in relative magnitude related to level of taxonomic aggregation?

2. Methods

2.1. Northeast lakes survey

A representative sample of the population of lakes in the NE was surveyed during summer months (July and August) from 1991 to 1996. The lakes were selected using a statistical survey design (Larsen *et al*., 1994; Stevens, 1994) so that the condition of the regional population of approximately 11,000 lakes could be inferred from the probability sample. The survey was designed to estimate the variance components described earlier.

At each lake in the survey, an array of physical, chemical, and biological attributes was measured, on which we examined the variability structure (attributes are identified in Figures 1 and 2). A complete description of these attributes and the biological sampling methods can be found in Baker *et al.* (1997) and Stemberger *et al.* (2001). The chemical indicators were organized into conservative and non-conservative indicators (Horne and Goldman, 1994; Wetzel, 2001). Conservative chemical attributes are the primary components of the ionic strength of lake water and include the major anions and cations. They are relatively inactive and should exhibit lowest temporal variability. Nonconservative chemical attributes are more reactive in lakes and should exhibit greater temporal variability. Trophic condition indicators capture lake primary production and may or may not exhibit more temporal variability than nonconservative chemical indicators. Two trophic condition indicators (Secchi depth and total phosphorus) were both measured directly and inferred from sediment diatom data. The zooplankton and fish assemblage indicators were organized into two categories: the number of individuals and the number of species (species richness). For zooplankton three levels of taxonomic organization were examined for both the individuals and richness attributes. All zooplankton species (the top level) were divided into rotifer species and crustacean species to form the second level. A third level was formed by subdividing crustacean species into calanoid, cladoceran, and cyclopoid species. An analogous, albeit less rigorous, taxonomic hierarchy was used for fish. A top level composed of all fish species was divided into minnow species, predator species, and other species to form the second level. For the third level predator species was subdivided into Centrarchid species and other predators, and other species was subdivided into water column species and benthic species. We evaluated whether temporal variation decreases as taxonomy is aggregated as suggested in Frost *et al.* (1992) and Kratz *et al.* (1994, 1995).

2.2. VARIANCE MODEL

A linear model that describes the data for the Northeast Lakes Survey is:

$$Y_{ijk} = \mu + L_i + T_j + LT_{ij} + I_{ijk} \ ,$$

where Y_{ijk} is the response for the kth visit at lake i during year j, μ is the overall mean, L_i is the random effect due to lake i, T_j is the random effect due to year j, LT_{ij} is the random effect due to the interaction of lake i and year j, and I_{ijk} is the random effect due to residual variation for the kth visit at lake i during year j. In this model subscript i ranges from 1 to l (the number of lakes in the survey), subscript j ranges from 1 to t (the number of years of data), and subscript k ranges from 0 to r_{ij} (the number of visits during year j at lake i). Let $\mathrm{Var}(L_i) = \sigma^2_{\text{Lake}}$ for all i, $\mathrm{Var}(T_j) = \sigma^2_{\text{Year}}$ for all j, $\mathrm{Var}(LT_{ij}) = \sigma^2_{\text{Interaction}}$ for all i and j, and $\mathrm{Var}(I_{ijk}) = \sigma^2_{\text{Residual}}$ for all i, j, and k. Total variance then consists of the sum of the four components: lake variance (σ^2_{Lake}), year variance (σ^2_{Year}), interaction variance ($\sigma^2_{\text{Iteraction}}$), and residual variance ($\sigma^2_{\text{Residual}}$). In terms of the variance model, status is the frequency distribution of the sum of the overall mean and the random effect due to lake i (i.e., $\mu + L_i$) across the lakes in the survey.

The variance model assumes that linear trend is not present. For the indicators we examined, there was no evidence for linear trend during the time period studied. When trend is present, its effect must be accounted for in the model; otherwise, trend will be lumped with the year variance component and result in biased estimation of that component. Linear trend is incorporated into the model by including a continuous variable composed of centered (mean zero) values based on the chronologic year data was collected. Two analyses then would be conducted: the first excluding the year class variable (random effect due to year) and the second including that variable. The first analysis provides a test for linear trend, and the second one provides a test for nonlinear effect due to years.

For each of the indicators, the variance model was fit to the full data set from the Northeast Lakes Survey. Estimates were obtained for the four variance components in that model: lake variance, year variance, interaction variance, and residual variance. Details regarding the estimation procedure are provided in the Appendix. A copy of the data used in this analysis can be obtained from D. P. Larsen.

2.3. Data transformations

Our survey design included revisits to some of the lakes during the summer sampling period (index window). Using the revisit data, we examined the relationship between the variance and the mean of each indicator by plotting the log of variance versus the log of mean, where variance and mean were calculated for each set of revisits. A tendency for variance to increase as the mean increased indicates nonhomogeneous residual variance for that indicator. Specifically, a slope of 2 for the log of variance versus the log of mean is indicative of an indicator that (a) has a constant coefficient of variation and (b) should be analyzed on the log scale. Based on this analysis, all of the chemical indicators except pH (which is measured on the log scale) and all of the trophic condition indicators except Secchi depth and diatom richness were analyzed on the log scale. For both zooplankton and fish assemblages, indicators for number of individuals were analyzed on the log scale, and indicators for species richness were analyzed on the original scale. Note that the linear model described previously applies to the scale on which measurements were analyzed, i.e., for transformed indicators the linear model applies to response values measured on the log scale.

3. Results

3.1. Variance component estimates

Variances were estimated for each of the four components and tested against zero (t-test for difference from zero), except residual variance. Estimates with p-values

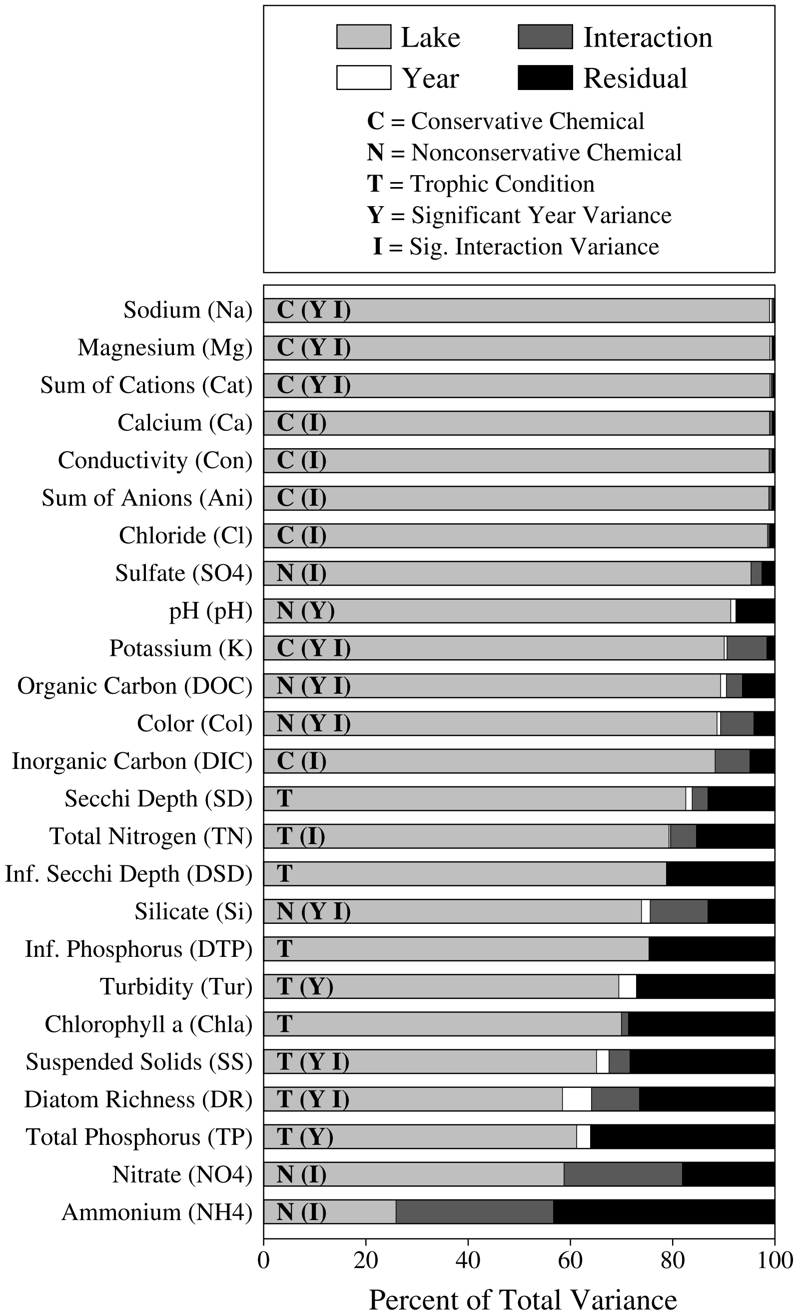

Figure 1. Estimated percent of total variance attributable to the lake, year, interaction, and residual components for chemical and trophic condition indicators ordered by size of residual plus interaction variance. The lake variance component was statistically significant for all indicators. Significance of year and interaction variance components is noted in the bar chart. By assumption for the variance model, residual variance is always greater than zero. Acronyms for indicator and indicator group are used in subsequent figures.

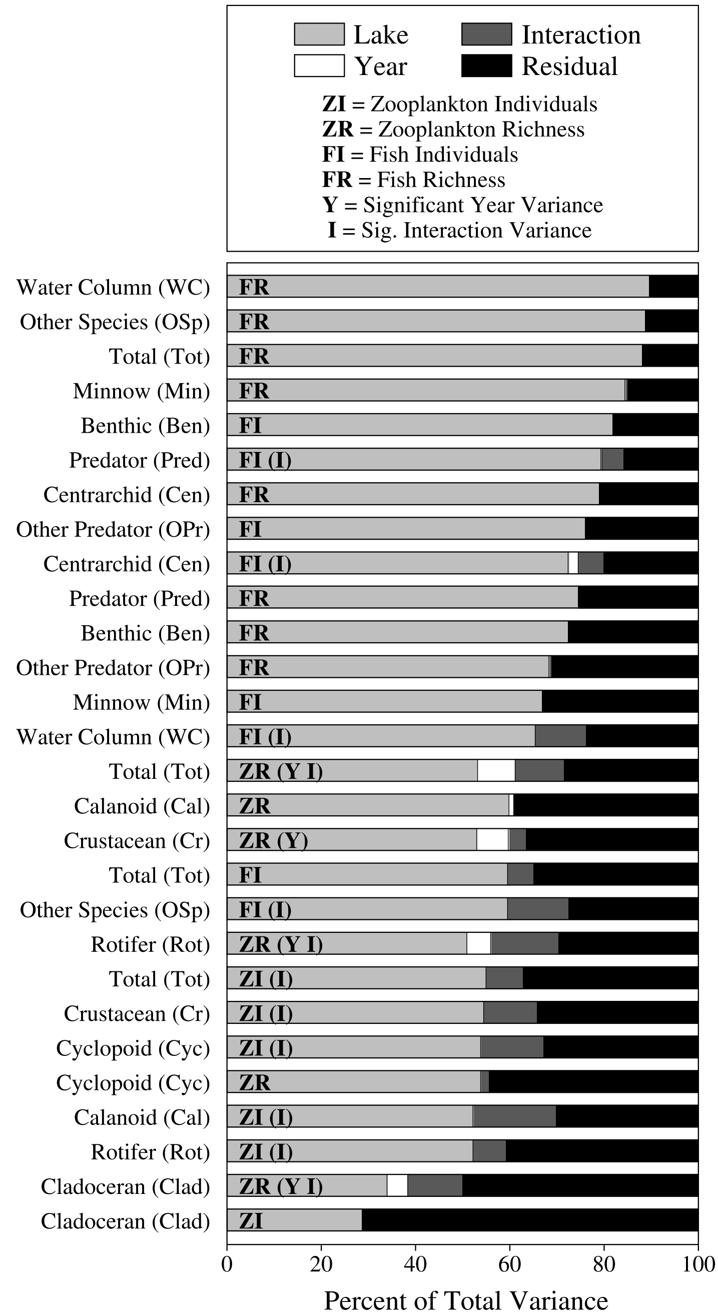

Figure 2. Estimated percent of total variance attributable to the lake, year, interaction, and residual components for zooplankton and fish indicators ordered by size of residual plus interaction variance. The lake variance component was statistically significant for all indicators. Significance of year and interaction variance components is noted in the bar chart. By assumption for the variance model, residual variance is always greater than zero. Acronyms for indicator and indicator group are used in subsequent figures.

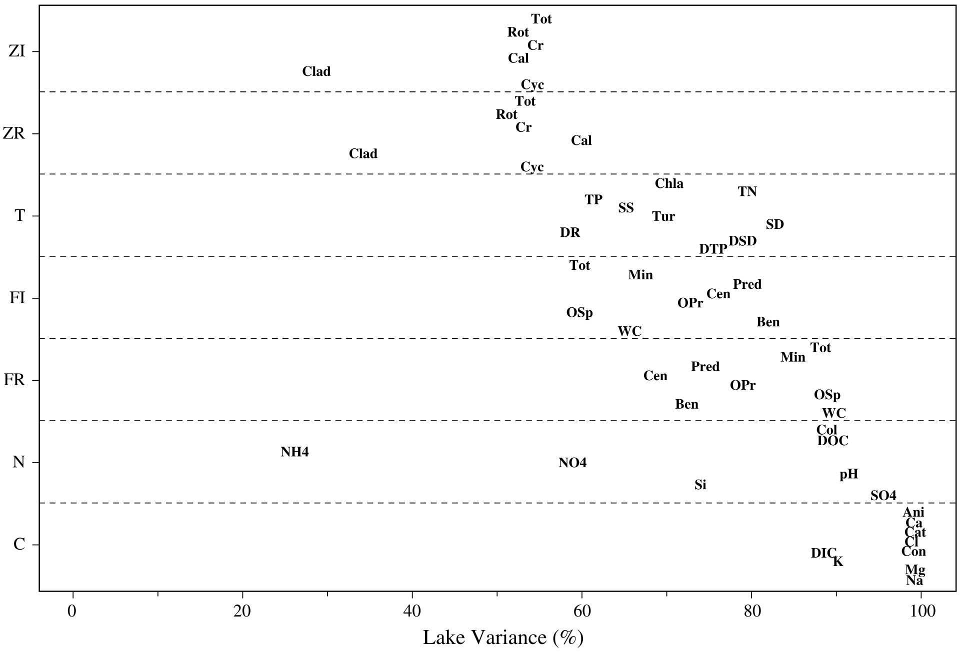

Figure 3. Percent of total variance attributable to the lake component, organized by indicator groups. Acronyms are decoded in Figures 1 and 2.

less than or equal to 0.1 were considered different from zero. By assumption for the variance model, residual variance is always greater than zero. For all indicators lake variance was always greater than zero with p-values typically no larger than 0.0001. The year and interaction variances exhibited greater ranges of p-values and in many cases were not significantly different from zero. The variance estimates are summarized as percent of total variance for each indicator (Figures 1 and 2). Within each indicator group, indicators are listed in order of increasing level of the sum of the interaction and residual variance proportions. These graphs illustrate the relative magnitude of the variance components and provide a quick comparison across the indicators with respect to bias in status estimation. To clarify some of the indicator patterns, indicator groups are also ordered by the percent of total variance attributable to the lake component (Figure 3); and the percent of total variance attributable to interaction variance is plotted against the percent attributable to residual variance (Figure 4). The general patterns in variance structure are summarized in the following sections.

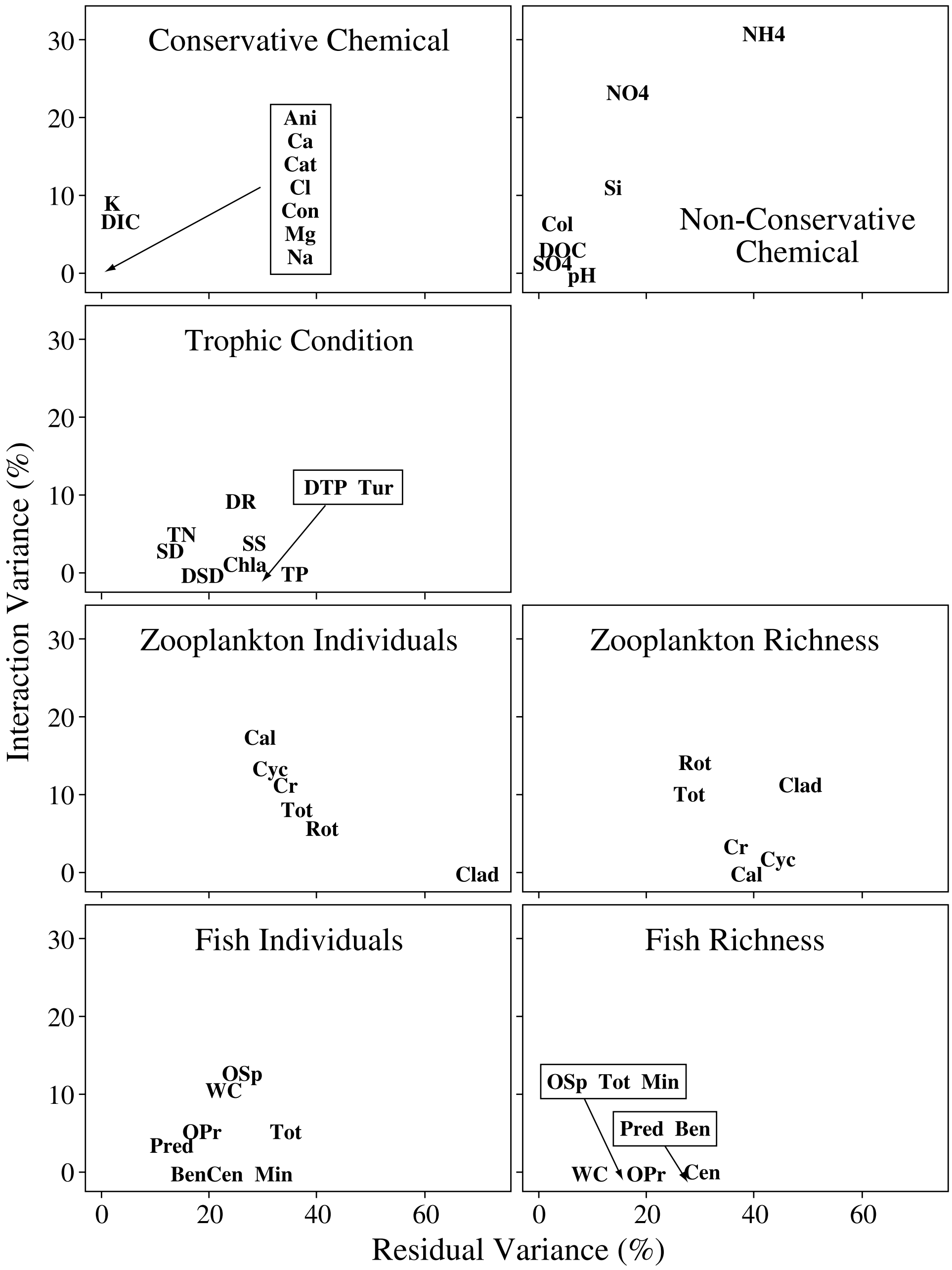

Figure 4. Percent of total variance attributable to interaction plotted against percent of total variance attributable to residual variance for the array of indicators in Figures 1 and 2. Acronyms are decoded in Figures 1 and 2.

3.2. Patterns in the structure of variation

3.2.1. *Lake Variance*

For the conservative and some nonconservative chemical indicators, lake variance accounted for substantially all of the variance observed (Figures 1 and 3). Two of the zooplankton indicators (cladoceran species and individuals) and the nutrient ammonium were at the other end of the spectrum, for which the lake component accounted for 35% or less of total variance (Figures 1–3). After the conservative and some nonconservative chemical indicators, fish richness indicators tended to have the highest percent of total variance attributable to the lake component (68–89%) with fish individuals intermediate (59–82%; Figures 2 and 3). Trophic indicators spanned a broad spectrum, from a high value for Secchi depth (about 85% attributable to lake variance) to a low value of about 58% for diatom richness (Figures 1 and 3). For the zooplankton indicators (both species richness and individuals), the percent of total variance attributable to lake was 51 to 60% (Figures 2 and 3) with a surprising consistency among both the individuals and richness indicators.

3.2.2. *Year Variance*

Indicators for which year variance was significantly different from zero are noted in Figures 1 and 2. For the indicators we examined, year variance accounted for none or a small percentage (less than 8%) of total variance. Year variance was zero for three groups: zooplankton individuals, fish individuals, and fish richness. For those indicators with nonzero year variance, the magnitude is inconsequential in the context of its influence on status estimation. However, see Urquhart *et al.* (1998) and Urquhart and Kincaid (1999) for insight into the effect of year variance on detection of trends.

3.2.3. *Interaction and Residual Variance*

Across all indicators, a consistent pattern emerged, with few exceptions, in which residual variance accounted for more (in some cases, substantially more) of the total variance than did interaction variance (Figure 4). That is, variability from all sources within the sampling index window (residual variance) was consistently higher than variability among years. For all of the conservative chemical indicators and four of the nonconservative chemical indicators, neither residual variance nor interaction variance accounted for much of total variance, so their magnitude compared to each other is of little concern. For the other three nonconservative chemical indicators (which include the biologically active nutrients ammonium and nitrate) both residual variance (13–43%) and interaction variance (11–31%) exceeded 10% of total variance. Residual variance for trophic condition indicators ranged from 13 to 29%, but interaction variance for those indicators was less than 10% of total variance. Residual variance for indicators of zooplankton individuals (30–71%) and zooplankton species richness (29–50%) accounted for a substantial part of total variance, whereas interaction variance exceeded 10% of total variance for half of the zooplankton indicators. For fish indicators, residual variance ranged from 16 to 35% for individuals and from 11 to 31% for species richness, whereas interaction variance exceeded 10% of total variance for only two fish individuals indicators and was close to zero for all of the fish species richness indicators.

Indicators for which interaction variance was significantly different from zero are noted in Figures 1 and 2. All of the conservative chemical indicators, all of the nonconservative chemical indicators except pH, but only three of the trophic condition indicators had significant interaction variance. For zooplankton, all of the indicators for individuals except cladocerans and half of the species richness indicators had estimates of interaction variance that were significantly different from zero. Four of the indicators for fish individuals but none of the fish species richness indicators were significantly different from zero.

3.2.4. *Taxonomic Hierarchy*

The effect of our hierarchical groupings on zooplankton and fish indicators was inconsistent (Figure 5). For indicators of zooplankton individuals and species richness, the percent of total variance attributable to the lake variance component ranged between 50 and 60% regardless of taxonomic level with the exception that cladocerans did not follow this pattern. Conversely, for indicators of fish individuals, there was a tendency for the percent of total variance attributable to the lake component to increase at finer levels of taxonomy; whereas no clear pattern emerged for fish species richness indicators.

With respect to the percent of total variance attributable to the residual component, proportional variance for indicators of zooplankton individuals tended to decline slightly, except for cladocerans, which increased substantially. For zooplankton species richness indicators, however, a clear pattern of increase with finer levels of the hierarchy occurred. For indicators of fish individuals, there was a pattern of decline with finer levels of the hierarchy but a tendency for variance to increase for fish species richness indicators. For fish, since the other variance components (year and interaction) typically accounted for a small percentage of total variance, the lake and residual variance components provided ‘mirror images’.

4. Discussion

4.1. Influence of Extraneous Variation on Status Estimation

The term ‘status’ is often used in a generic way to convey a snapshot of the condition of a resource at an instant in time, a slice through time. For many resource types, however, obtaining a slice through time is not feasible, and instead the snapshot reflects an average during a specified time period, an index window. In this case, each visit to a sample site and the measurements taken during that visit can be considered an estimate of the average for that sample site during the index

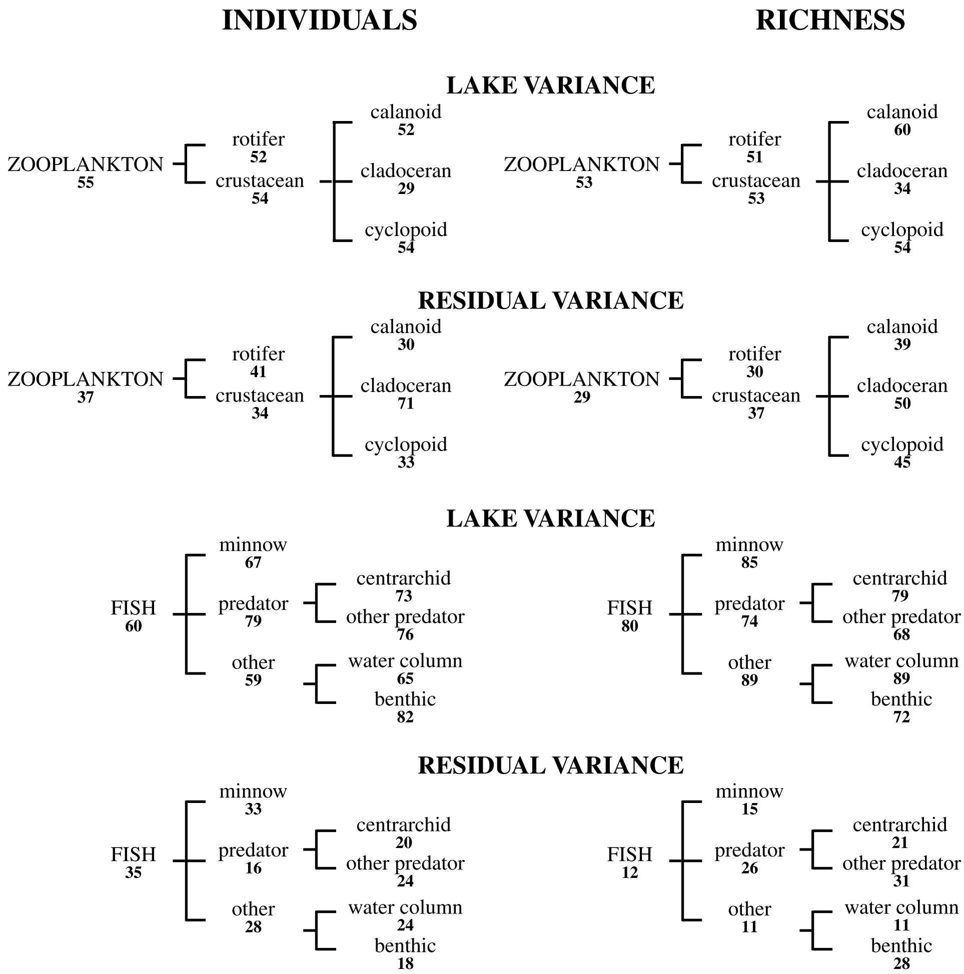

Figure 5. Percent of total variance attributable to the lake or residual components as a function of taxonomic hierarchy for the zooplankton and fish individuals and richness indicators.

window. One expression of status of the target population is the frequency distribution of the set of average scores for an attribute of interest across all elements of the population, summarized as the population cumulative distribution function (CDF), e.g., Figure 6. Well designed sample surveys allow us to infer the population's CDF by obtaining a set of measurements on a properly selected sample of the population elements. However, a variety of 'errors' arise that interfere with population description or inference.

Overton (1989) characterizes the set of nonsampling variable errors as 'extraneous' and, with the exception of rounding errors, considers them 'measurement errors'. This concept of measurement errors differs from analytical measurement errors estimated by replicate measurements in that it includes some temporal and

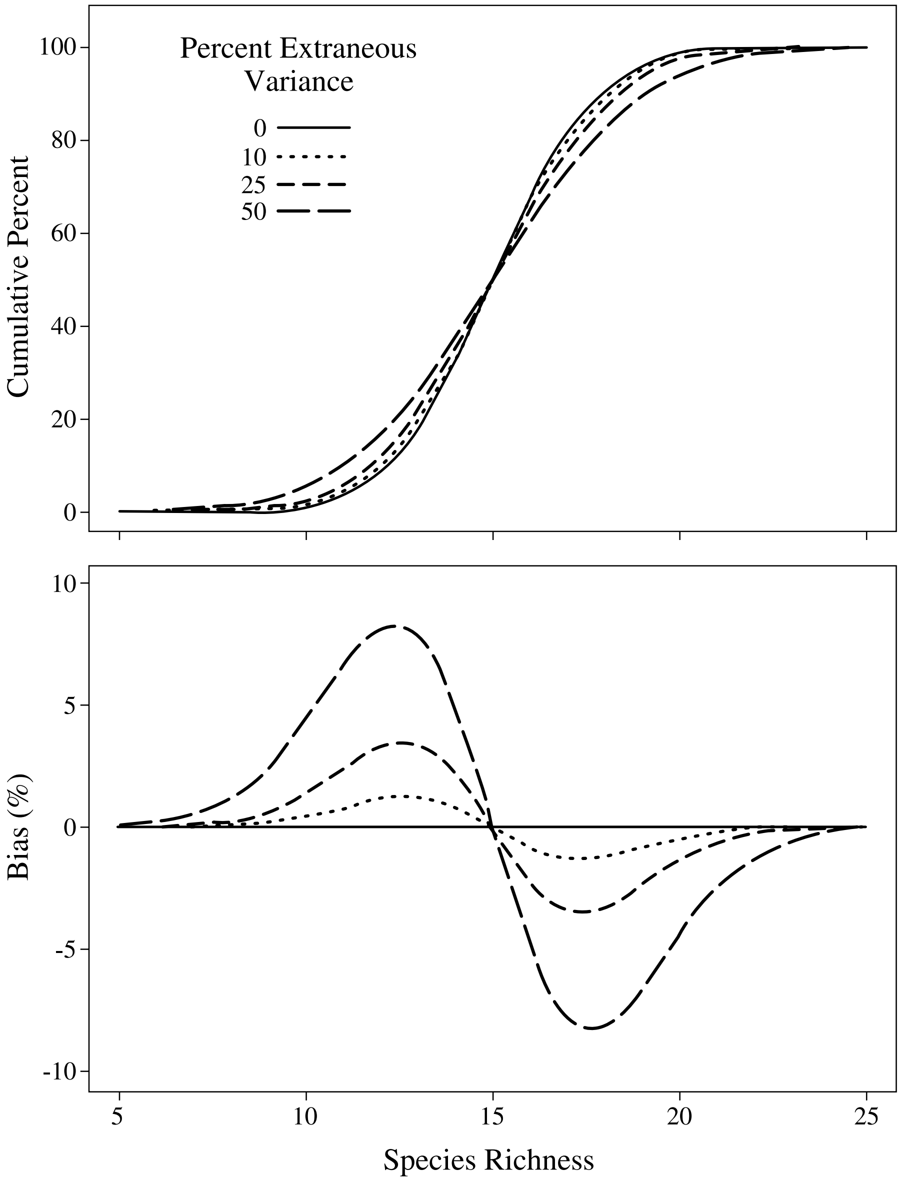

Figure 6. (Top) The bias introduced into status descriptions (cumulative distribution functions) by extraneous variance for a Normal distribution to which various levels of extraneous variance have been introduced. Bias is 0 at the mean for Normal distributions, and maximum at $\pm$ one standard deviation from the mean. The true distribution (solid line) is spread further and further as extraneous variance increases. (Bottom) The pattern of bias introduced by extraneous variation is clear in this panel.

spatial components. Collectively, these measurement errors create bias in the CDF. Other authors merely call these errors measurement errors or replication errors. In this paper we use the term extraneous variance to reference these 'measurement errors'. We classify residual variance and interaction variance as extraneous variance and evaluate the impact of this variance on estimation of status of the population.

Overton (1989) evaluated the relationship between the magnitude of extraneous variance and bias in the CDF. Extraneous variance spreads the CDF resulting in biased estimation of status (Figure 6). In the top panel of Figure 6 the Normal distribution is used to simulate the CDF for species richness of, say, fish in a finite population of lakes. The solid line is the CDF (i.e., the status description) in the absence of extraneous variance, which is the true status description of the population. The dashed lines are the CDFs in the presence of extraneous variance, where extraneous variance is expressed as percent of total variance. As the amount of extraneous variance increases, bias in some parts of the CDF increases (see the following discussion). For example, extraneous variance of 10% results in a slight overestimate of the percent of the resource less than 12 (10% versus 9%), whereas extraneous variance of 50% results in a large overestimate of that percentage (17%). For a given level of extraneous variance, however, bias is not constant across the CDF. At the median, for example, the CDF is not affected by extraneous variance. For the Normal distribution, bias reaches a maximum at ± 1 standard deviation from the mean, illustrated in lower panel of Figure 6. For asymmetrical distributions, Overton (1989) noted that small levels of extraneous variance can still produce significant bias in the CDF. Specifically, he stated that, for a population with an exponential distribution, there is the potential for a small level of extraneous variance to produce substantial bias in the lower tail of the CDF.

The discussion above centers around a description of status within a particular year. However, as referenced in the description of the variance model, it is also reasonable to refer to status as the frequency distribution of fundamental differences among lakes, in which case the CDF contains biases introduced by the residual and interaction components of variation. Taken together, these sources of variation confound the description of the true distribution of differences among lakes.

Performance of the indicator groups regarding status estimation is illustrated in Figure 3, which displays proportion of the lake component of variance. The conservative chemical indicators performed best with respect to status estimation. Excluding dissolved inorganic carbon and potassium, which exhibited slighted lower levels of lake variance (in the range 88–90%), the conservative chemical indicators had lake variance of 98% or higher. The nonconservative chemical indicators exhibited a wide range of lake variance. Four of those indicators (color, organic carbon, pH, and sulfate) performed similarly to dissolved inorganic carbon and potassium, but the nutrient indicators (ammonium and nitrate) had low levels of lake variance. After the conservative chemical indicators and the nonconservative chemical indicators listed previously, the fish indicators performed best. Among fish indicators, the richness indicators were slightly better than the

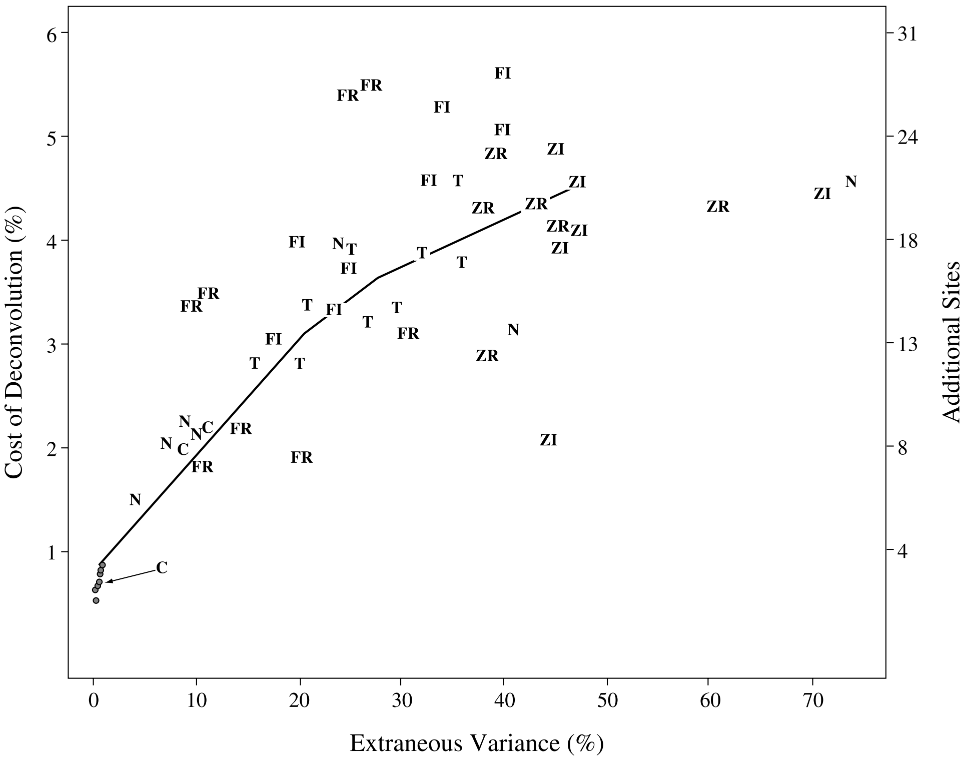

Figure 7. The cost of deconvolution (correcting a CDF for biases introduced by extraneous variance) can be expressed as the average increase in width of the confidence bounds of the CDF, here expressed as a function of the magnitude of extraneous variance for the array of lake indicators summarized in this study. Cost can also be expressed in terms of the additional sampling necessary to achieve the same precision if extraneous variance were 0, here illustrated for a case of a random sample of 50 lakes from a large population of lakes. A robust regression line has been fit to the points to illustrate the general pattern. Indicators with extraneous variance exceeding 50% (total of three indicators) were excluded from calculation of the regression line. Acronyms are decoded in Figures 1 and 2.

individuals indicators. Performance of the trophic condition indicators was similar to the fish individuals indicators. Zooplankton indicators performed the worst in terms of status estimation. Lake variance for zooplankton indicators (both richness and individuals indicators) occurred within a narrow range except for cladoceran species, which exhibited substantially lower levels of lake variance.

When the residual and interaction components of variation are sufficiently large to introduce bias in the estimated CDF, that bias can be removed by incorporating the estimated residual and interaction variation into a process called deconvolution (e.g., Gaffy, 1959). However, there is a cost associated with deconvolution: an increase in the sampling uncertainty of population inferences from a sample. Sampling uncertainty directly affects the precision with which a quantity (in this case the CDF) is estimated, i.e., increased sampling uncertainty means decreased precision. The increase in uncertainty of population inferences resulting from deconvolution is illustrated in Figure 7, which plots cost of deconvolution against percent extraneous variance for each of the lake indicators. Cost is assessed two ways: (1) the average increase in width of the confidence bounds of the CDF and (2) the approximate increase in sample size needed to achieve the same sampling precision that would be achieved if extraneous variance were 0. To illustrate the first type of cost, suppose the confidence bound at a chosen value was 20±5% and the cost was 4%, then deconvolution increased width of the confidence bound from 10 to 14%, producing a confidence bound of 20±7%. The second type of cost is illustrated in Figure 7 under the assumption that a random sample of 50 lakes has been selected from a large population of lakes. A clear pattern of increasing cost as the percent of extraneous variance increases is evident. A robust regression line is included in Figure 7 to aid in visualizing the overall pattern. Due to the lack of indicators with extraneous variance in excess of 50%, three indicators with high levels of extraneous variance were excluded from calculation of the regression line.

Year variance affects the status estimate in a slightly different way. Whereas the effect of extraneous variation is to 'rotate the CDF in a pinwheel fashion around the mean for Normal distributions, year variance shifts the entire CDF a constant amount. The size of each year's shift is dependent on the magnitude of the year effect (see the variance model). Bias introduced by year variance thus can be removed by estimating the year effect for a particular year (e.g., Searle *et al.*, 1992) and subtracting that value from the CDF estimate. As noted in the Results Section, the year effect typically was minor for the indicators we examined.

4.2. Interpretation of Residual and Interaction Variance

Although careful attention to the selection of sites and the pattern of visits to sites within and across years is an important aspect of designing and modifying surveys (Urquhart *et al.*, 1998; Urquhart and Kincaid, 1999), equally important is evaluation of the structure of variation. An evaluation of the latter can be applied to improving sample collection and processing protocols with the goal to reduce bias. Of particular importance is residual variation: can its magnitude be reduced by changes in collection or measurement protocols? For example, the fish richness indicators are characterized by a moderate amount of residual variation. Within a particular lake, it is unlikely that the number of species resident during the index window changed; it is more likely that not all the species present were captured (Vaux *et al.*, 2000). Therefore, residual variation probably reflects the ability to capture the resident species when the lake is sampled. This is not an unreasonable assumption given the challenges of sampling fish in multiple habitats of lakes using multiple kinds of gear located in various parts of the lake. Thus, likely reduction of residual variation would come from improvement in the species collection methods.

Similarly, sediment diatom richness is characterized by moderate residual variation. Here again, it is unlikely that the composition of diatoms collected in the top centimeter of lake sediment would change substantially during the index window (Cumming *et al.*, 1992; Dixit *et al.*, 1999). The surface centimeter of sediment likely contains diatoms representing several years accumulation, so many new species are unlikely to settle during the index window of a particular year. The results presented here are based on a single diatom core near the deepest part of the lake (Dixit *et al.*, 1999). Residual variation estimated by taking cores in the same lake at different times during the index window likely represents spatial variation in the composition of sediment diatoms. Reduction in residual variation might result from sampling at several locations to better estimate an average condition during the index window.

In contrast, indicators like total phosphorus, chlorophyll a, nutrients, and zooplankton individuals, can vary substantially within the index window (compared to the fish and sediment diatom richness indicators). The consequent moderate to high residual variation summarized here could be managed by increased sampling during each visit to the lake but is more likely to be sensitive to increasing the number of times a lake is visited during the index window to obtain a more precise estimate of the mean during this period. Of course, the improvement derived by revisiting lakes must be balanced by the cost of not visiting other lakes. See Urquhart *et al.* (1998) for an examination of the tradeoffs between revisiting sites or visiting additional sites in a survey. Stemberger *et al.* (2001) also describe how spatial partitioning of the lake population (for example, by ecological regions) can be useful for managing some of the components of variation.

Interaction variance is a natural feature of these systems, and as such is not subject to management by changes in sampling protocols as residual variation might be. Revisiting a sufficient number of lakes to estimate its magnitude, then using a process like deconvolution to remove its effects on status estimation is the primary option for managing this component of variation with respect to status estimation. For most of the lake indicators described here, this component of variance is relatively small compared with residual variation (to a lesser degree) and lake variation (to a greater degree), so the cost of removing its effect would be relatively low.

4.3. Variance patterns

The patterns of variation across indicator groups are what is expected from an evaluation of biological or geochemical activity of the indicators (Figures 1 and 2). Almost all of the variation among the conservative chemical indicators is associated with lake-to-lake differences; very little variation is nonspatial. A few of the nonconservative chemical indicators (color, organic carbon, pH, and sulfate) have nearly the same pattern but with slightly increased interaction and residual variation. On the other hand, nutrients (ammonium and nitrate) can be expected to be quite variable, as they are consumed and recycled rapidly during the summer sampling window and can be affected from year to year by local climate. Consequently, the nutrient indicators exhibit high percentages of total variation that are residual and interaction.

Of the trophic condition indicators, lake transparency measured with a Secchi disk contains the highest percent of total variance associated with the spatial component. Therefore, from the perspective of estimating the trophic status of lakes, this indicator would provide the least biased estimation, provided that transparency in the set of lakes under consideration was controlled by algal growth and not by nonalgal fine sediments or other factors. This is perhaps fortuitous because its use is inexpensive and the protocol can easily be taught to nonprofessionals (Klang, 1997; Lee *et al.*, 1997). In fact, extensive volunteer networks use this tool to evaluate lake trophic condition in north-temperate lakes, and relationships between this indicator and other trophic condition indicators are well documented (Carlson, 1977).

Frost *et al.* (1992) and Kratz *et al.* (1994, 1995) indicate that variability can be associated with level of taxonomic aggregation, i.e., species level is more variable than guild level which is more variable than major groupings. However, our results indicate that aggregation produces no clear pattern with respect to proportioning total variation. For example, for zooplankton at the highest level of aggregation, nearly the same percent of total variation is attributable to the lake component for both richness and individuals indicators. The number of individuals is clearly more variable than the much more aggregated number of species; however, on a relative scale, the decomposition of total variation is nearly the same for both individuals and richness. For fish, patterns also are inconsistent. For example, splitting total fish richness into minnows, predators, and other species decreases the percent of total variation associated with spatial variation for predators but does not affect the percent for minnows or other species. Splitting predators to a finer level (centrarchids and other predators) makes little difference.

Acknowledgements

Conducting regional ecological surveys requires the coordination and cooperation of numerous individuals responsible for planning, logistical support, field sampling, analytical processing, quality control and assurance, and data management. We gratefully acknowledge those individuals too numerous to list by name. In addition, comments by Thomas Whittier, Richard Stemberger, Steve Stehman, and F. Jay Breidt improved the manuscript. The research described in this manuscript has been funded by the U.S. Environmental Protection Agency under contracts 68-C6-0005 and 68-D0-1005 with Dynamac International Corporation, cooperative agreement CR-824682 with the Department of Fisheries and Wildlife at Oregon State University, and cooperative agreement CR-829095 with the Department of

Statistics at Colorado State University. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

Appendix: Variance Component Estimation Procedure

The method of moments procedure was used to calculate variance component estimates. For a through discussion of that procedure, see Searle (1971) and Searle *et al.* (1992); a brief description follows. The procedure first is summarized for an equal probability sampling design and then is generalized to handle an unequal probability design. As an initial step, an analysis of variance was conducted. The analysis of variance provided an observed mean square value for each of the factors in the model. The next step was to calculate the expected mean square for each factor. The expected mean squares express each model factor as a linear combination of the variance components. The observed mean squares were then equated to the expected mean squares, and the resulting series of equations was solved to produce the variance component estimates. Note that many statistical programs, e.g., SAS, are capable of calculating the expected mean squares. We used a SAS Type III analysis to obtain the observed and expected mean squares.

Since the EMAP Northeast Lakes Survey was an unequal probability sampling design, we wanted to incorporate the sample weights into the estimation procedure. We accomplished that goal by including the inverse of the sample inclusion probabilities, i.e., the sample expansion factors, as weights in PROC GLM of SAS. The impact of doing a weighted analysis in PROC GLM is that all of the observed mean squares are inflated by the weights. For all of the expected mean squares except the one for residual error (our residual variance), the coefficients for the expected mean squares are inflated by the weights. Thus, for all variance components except the one for residual variance, valid estimates are obtained from the weighted analysis. For residual variance the single coefficient in the expected mean square is unaffected by the weights, and so a valid estimate is not obtained. Therefore, the following estimator was used to estimate residual variance:

$$\hat{\sigma}^2_{\text{Residual}} = \frac{1}{n-p}\left\{ n \frac{\sum_{i=1}^{n} w_i \left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{n} w_i} \right\},$$

where n is the sample size, p is the number of parameters in the model, w_i are the sample weights, y_i are the response values, and $\hat{y}_i$ are the predicted response values. For an equal probability design, this estimator reduces to the usual estimator for residual error.

References

Baker, J. R., Peck, D. V. and Sutton, D. W. (eds.): 1997, *Environmental Monitoring and Assessment Program – Surface Waters: Field Operations Manual for Lakes*, EPA/620/R-97/001, U.S. Environmental Protection Agency, Washington, D.C.

Carlson, R. E.: 1977: 'A trophic state index for lakes', *Limnol. Oceanogr.* **22**, 363–369.

Cumming, B. F., Smol, J. P., Kingston, J. C., Charles, D. F., Birks, H. J. B., Camburn, K. E., Dixit, S. S., Uutala, A. J. and Selle, A. R.: 1992, 'How much acidification has occurred in the Adirondack region (New York, U.S.A.) lakes since pre-industrial times?', *Can. J. Fish. Aquat. Sci.* **49**, 129–141.

Dixit, S. S., Smol, J. P., Charles, D. F., Hughes, R. M., Paulsen, S. G. and Collins, G. B.: 1999, 'Assessing water quality changes in the lakes of the northeastern United States using sediment diatoms', *Can. J. Fish. Aquat. Sci.* **56**, 131–152.

Frost, T. M., Carpenter, S. R. and Kratz, T. K.: 1992, 'Choosing Ecological Indicators: Effects of Taxonomic Aggregation on Sensitivity to Stress and Natural Variability', in: D. H. McKenzie, D. E. Hyatt and V. J. McDonald (eds), *Ecological Indicators*, Elsevier Applied Science, New York, pp. 215–227.

Gaffy, W.: 1959, 'A consistent estimator of a component of a convolution', *Ann. Math. Stat.* **30**, 198–205.

Horne, A. J. and Goldman, C. R.: 1994, *Limnology*, 2nd ed., McGraw-Hill, New York.

Klang, J. L. K.: 1997, *Citizen Lake-Monitoring Program: 1996 Report on the Transparency of Minnesota Lakes*, Minnesota Pollution Control Agency, Division of Water Quality, St. Paul, Minnesota.

Kratz, T. K., Magnuson, J. J., Frost, T. M., Benson, B. J. and Carpenter, S. R.: 1994, 'Landscape Position, Scaling, and the Spatial and Temporal Variability of Ecological Parameters: Considerations for Biological Monitoring', in: S. L. Loeb and A. Spacie (eds), *Biological Monitoring of Aquatic Ecosystems*, Lewis Publishers, Boca Raton, Florida, pp. 217–231.

Kratz, T. K., Magnuson, J. J., Bayley, P., Benson, B. J., Berish, C. W., Bledsoe, C. S., Blood, E. R., Bowser, C. J., Carpenter, S. R., Cunningham, G. L., Dahlgren, R. A., Frost, T. M., Halfpenny, J. C., Hansen, J. D., Heisey, D., Inouye, R. S., Kaufman, D. W., McKee, A. and Yarie, J.: 1995, 'Temporal and Spatial Variability as Neglected Ecosystem Properties: Lessons Learned from 12 American Ecosysems', in: D. Rapport and P. Calow (eds), *Evaluating and Monitoring the Health of Large-Scale Ecosystems*, Springer-Verlag, New York, pp. 359–383.

Larsen, D. P., Thornton, K. W., Urquhart, N. S. and Paulsen, S. G.: 1994, 'The role of sample surveys for monitoring the condition of the nation's lakes', *Environ. Monit. Assess.* **32**, 101–134.

Larsen, D. P., Urquhart. N. S. and Kugler, D. L.: 1995, 'Regional scale trend monitoring of indicators of tropic conditions of lakes', *Water Resour. Bull.* **31**, 117–140.

Larsen, D. P., Kincaid, T. M., Jacobs, S. E. and Urquhart, N. S.: 2001, 'Designs for evaluating local and regional scale trends', *BioScience* **51**, 1069–1078.

Lee, J., Binkley, M. R. and Carlson, R. E.: 1997, 'The great American Secchi dip-In', *GIS World* **10**, 42–44.

Lewis Jr., W. M.: 1978, 'Comparison of temporal and spatial variation in the zooplankton of a lake by means of variance components', *Ecology* **59**, 666–671.

Matthews, W. J.: 1990, 'Spatial and temporal variation in fishes of riffle habitats: A comparison of analytical approaches for the Roanoke River', *Am. Midl. Nat.* **124**, 31–45.

Messer, J. J., Linthurst, R. A. and Overton, W. S.: 1991, 'An EPA program for monitoring ecological status and trends', *Environ. Monit. Assess.* **17**, 67–78.

Olsen, A. R., Sedransk, J., Edwards, D., Gotway, C. A., Liggett, W., Rathbun, S., Reckhow, K. H. and Young, L. J.: 1999, 'Statistical issues for monitoring ecological and natural resources in the United States', *Environ. Monit. Assess.* **54**, 1–45.

Overton, W. S.: 1989, 'Effects of Measurement and Other Extraneous Errors on Estimated Distribution Functions in the National Surface Water Surveys', *Technical Report 129*, Department of Statistics, Oregon State University, Corvallis, Oregon.

Overton, W. S., White, D. and Stevens, D. L.: 1991, *Design Report for EMAP, the Environmental Monitoring and Assessment Program*, EPA/600/3-91/053, U.S. Environmental Protection Agency, Washington, D.C.

Searle, S. R.: 1971, *Linear Models*, John Wiley & Sons, New York.

Searle, S. R., Casella, G. and McCulloch, C. E.: 1992, *Variance Components*, John Wiley & Sons, New York.

Stemberger, R. S., Larsen, D. P. and Kincaid, T. M.: 2001, 'Sensitivity of zooplankton for regional lake monitoring', *Can. J. Fish. Aquat. Sci.* **58**, 2222–2232.

Stevens, D. L.: 1994, 'Implementation of a national monitoring program', *J. Environ. Manage.* **42**, 1–29.

Urquhart, N. S., Paulsen, S. G. and Larsen, D. L.: 1998, 'Monitoring for policy-relevant regional trends over time', *Ecol. Appl.* **8**, 246–257.

Urquhart, N. S. and Kincaid, T. M.: 1999, 'Designs for detecting trend from repeated surveys of ecological resources', *J. Agric. Biol. Envir. S.* **4**, 404–414.

Vaux, P. D., Whittier, T. R., DeCeasre, G. and Kurtenbach, J. P.: 2000, 'Evaluation of a backpack electrofishing unit for multiple lake surveys of fish assemblage structure', *N. Am. J. Fish. Manage.* **20**, 168–179.

Wetzel, R. G.: 2001, *Limnology: Lake and River Ecosystems*, 3rd ed., Academic Press, New York.